DOCUMENT RESUME

ED 452 070 SE 064 671

AUTHOR MacIsaac, Dan; Sawada, Daiyo; Falconer, Kathleen
TITLE Using the Reformed Teaching Observation Protocol (RTOP) as a Catalyst for Self-Reflective Change in Secondary Science Teaching.
PUB DATE 2001-04-13
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Change; *Reflective Teaching; *Science Instruction; *Science Teachers; Secondary Education; Teaching Methods
IDENTIFIERS *Reform Strategies

ABSTRACT

This report examines the feasibility of using a measurement instrument, the Reformed Teaching Observation Protocol (RTOP), to catalyze teacher self-reflective change. RTOP was used in a summer course as a tool to help understand and even operationally define reform teaching. RTOP assesses four major pedagogical domains: (1) lesson design and implementation; (2) propositional and procedural knowledge; (3) communicative interactions; and (4) student-teacher relationships. Of the current instruments, RTOP is the only one satisfying the following criteria: (1) focused on math and science; (2) aimed at classroom K-20; (3) focused exclusively on reform rather than general characteristics such as classroom management, lesson closure, etc.; (4) brief to administer; (5) very high inter-rater reliability; (6) factor analyzed for construct validity; (7) proven predictive validity; and (8) readily available training and reference manuals. (Contains 41 references.) (SAH)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 452 070

Using the Reformed Teaching Observation Protocol (RTOP) as a Catalyst for Self-Reflective Change in Secondary Science Teaching

by
Dan MacIsaac
Daiyo Sawada
Kathleen Falconer

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Dan MacIsaac

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

SE064671

Using The Reformed Teaching Observation Protocol (RTOP) As A Catalyst For Self-Reflective Change In Secondary Science Teaching

Teacher Self-Reflective Change – Empowered by RTOP

Dan MacIsaac, Department of Physics & Astronomy, Northern Arizona University, Campus Box 6010, Flagstaff AZ 86011-6010;
danmac@nau.edu
Daiyo Sawada, University of Alberta
Kathleen Falconer, Arizona State University

In recent years, a number of classroom observation instruments have been developed to provide both qualitative and quantitative data to document and describe science and mathematics teaching in this era of standards-based reform (e.g., Calhoun et al, 1997; Forman, 1996; Horizon Research Inc., 1998; Lai & Young, 1997; Lawson, 1995; Lipsey, 1994; Mayer, 1999: Newmann, Marks & Gamoran, 1995; O'Sullivan, 1995; Piburn, Sawada, Turley, Falconer, Benford, Bloom & Judson, 2000; Porter, 1995; Priestly et al, 1998; QUASAR, 1992; Sawada, Piburn, Turley, Falconer, Benford, Bloom, & Judson, 2000; Silver & Cai, 1993; Tittle & Pape, 1996; Tomoff, 2000; Weir, 2000; Young, Brett, Squires & Lemire, 1995). These instruments provide data that is often used in the evaluation of major projects designed to prepare reformed teachers of mathematics and science. Typically, data from these instruments allow evaluators to determine whether reformed pre-service or inservice programs produce teachers who teach in a more reformed manner than teachers not participating in such programs.

However, in addition to assessing the products of reform, such instruments can also be used to catalyze change by engaging teachers in self-reflection on their own practice. The present paper describes a study that used such an instrument as a catalyst for self-reflective change toward reformed teaching of mathematics and science. The instrument of choice was the Reformed Teaching Observation Protocol (RTOP) (Sawada et. al. 2000; Piburn et. al. 2000). Based on standards-based inquiry (AAAS, 1989; NCTM, 1989, 1991, 1995, 2000; NRC, 1996, 2000), RTOP assesses five major pedagogical domains: (1) lesson design and implementation, (2) propositional and procedural knowledge, (3) communicative interactions and (4) student teacher relationships. Of the current instruments available, RTOP is the only instrument satisfying all of the following criteria: (1) focused on both mathematics and science, (2) aimed at classrooms from K to 20, (3) focused exclusively on reform rather than general characteristics such as classroom management, lesson closure, etc., (4) brief to administer, (5) very high inter-rater reliability (r = 0.90 and above), (6) factor analyzed for construct validity, (7) proven predictive validity (correlations with student achievement gains are typically in the 0.70 – 0.95 range), and (8) training and reference

manuals are available. We consider RTOP to be an operational definition of reform teaching practice for the remainder of this paper.

Other researchers have developed rubrics, matrices, models, inventories, conceptual grids and other frameworks to characterize reform. In a particularly relevant example, Adams and Krockover (1997, 1999) have described the use of the Secondary Science Teacher Analysis Matrix (STAM) from the Salish project (Gallagher & Parker, 1995) to guide teacher development during the early years of their profession. Adams and Krockover (1997, 1999) described how they made use of STAM codings to characterize classroom behaviors along a six category continuum as *didactic, transitional, conceptual, early constructivist, experienced constructivist and constructivist inquiry* (Gallagher & Parker, 1995). As part of a two year case study, one teacher was provided with a copy of the STAM characteristic descriptor matrix. The following fall, 'his teaching style dramatically changed' as he deliberately sought to change his STAM ratings, using the STAM matrix descriptors to stimulate recall of experiences and events from his pre-service teacher training which he then deliberately incorporated into his teaching. Adams and Krockover (1997, 1999) concluded that models such as STAM could be used to stimulate situationally-appropriate recall of pre-service teacher preparation, 'helping teachers to refocus on research-based instructional methods'. They described STAM as a heuristic device for guiding their subject's development as a teacher. Priestly, Priestly, Sutman, Schmuckler, Hilosky & White (1998) describe similar usages for their "Inquiry Matrix".

Focus of this Study

Given the successful use of conceptual grids for supporting teacher change, the present study examined the feasibility of using a measurement instrument, the RTOP, to catalyze teacher self-reflective change. During a summer course we deliberately sought to foster teacher change by explicitly introducing the RTOP as a tool to help understand and even operationally define (Arons, 1997) reformed teaching.

The Setting & Participants

During the summer of 2000, a three week graduate credit course was conducted for inservice and pre-service teachers at a midsized western state university specializing in teacher education. The graduate credit course was underwritten with grant funds awarded through the Federal Eisenhower Mathematics and Science Education Act, whose intent is to improve teacher skills and quality of instruction in mathematics and science (ABOR, 2000; MacIsaac & Cole, 1999). There were thirty-seven participants, including three master-teacher instructors, two science education faculty, four part-time support staff who were pre-service secondary science teachers and twenty-eight teachers enrolled for credit. Of the twenty-eight registrants, twenty five were inservice teachers (twenty-one from public schools and four from charter schools) and three were pre-service teachers. Two of the registrants were from out of state schools, and thirteen were from the single surrounding school district, with two more from nearby charter schools. All participants received stipends of varying amounts.

By grade, nineteen of twenty-eight registrants were high school teachers, with six primarily middle school and the remainder mixed grade levels. By subject, most of the middle school teacher registrants (four of six) were primarily science teachers who also taught math, the other two were the reverse of this mix. High school teacher registrants were mostly mixed science subject specialists (mixed science and chemistry for eight; three were principally physics, two biology and one earth science -- all taught several subject areas) while eight identified themselves as primarily math teachers (only two usually taught nothing but math).

Data Collection

Several sources of data were collected routinely throughout the course. These consisted of autobiographical material, journal reflections, videotaped lessons, audiotaped focus group discussions, survey questionnaires, RTOP scores on participants and instructors and unit plans.

Chronology of The Course

The course was designed to bring together a small number of high school and middle school science and mathematics teachers and foster the preparation of classroom activities (lesson plans, unit plans, curricula, activities and materials) that met the recently issued state standards for science and mathematics content. The philosophical and instructional guides for this course were taken from research-based ideas featured in the high school modeling physics curriculum (Wells, Hestenes & Swackhamer, 1995; Hallouin & Hestenes, 1987; Hestenes, 1987). Modeling pedagogy incorporates extensive model-building, electronic probe data acquisition and computer analysis of data, data interpretation, negotiation of meaning, collaborative learning and peer presentation (MacIsaac & Cole, 1999).

In planning the course, the initial thought was to use the RTOP as a pretest-posttest evaluation tool – the teaching of participants would be assessed early in the course and then again later in the course to measure the change. However, such usage would be very "unreformed," rather like using a yardstick to measure when those being measured have no idea what a yard is. The alternative of using the RTOP as an inquiry tool took hold. In the hands of participants it could be used to acquire data about reformed teaching, their own teaching and the teaching of others. The decision was made to bring the RTOP into the course as an inquiry tool available to all to improve their teaching through inquiry. With this change in plans, the focus of the research changed to an examination of the feasibility of using RTOP as a catalyst for inquiry teaching.

Week One

With this new role for RTOP, the summer course began with an RTOP workshop. All participants were provided with copies of the instrument for reading and discussion. The background history of the instrument and its use in other studies (Piburn, Sawada,

Turley, Falconer, Benford, & Bloom, 2000) were presented by one of the authors of the instrument. Next, a fifteen minute videoclip of a mathematics lessons was viewed and participants were asked to rate the instructional experience using the RTOP instrument. The taped selection was deliberately selected to show seductively interesting and "attractive" instruction that nonetheless receives low scores from trained observers. Participants rated the instruction privately then reported their scores aloud. Participants were then asked to form groups of three such that each group contained members who had scored the lesson relatively low, moderate and relatively high. Groups the discussed, debated and revised their scores, which usually fell drastically (E.g. from RTOP scores of 80% to 20%). In the process, an appreciation and understanding of the nature and intents of RTOP began to emerge.

Next all participants watched a second short video of a modeling physics (Hestenes, 1987) lesson which trained RTOP observers had rated very highly. Again, teachers rated the video individually and reported their numeric scores aloud. A whole group discussion of the second tape followed, together with whole group questioning of the RTOP author. Questions ranged from the validity of the instrument for elementary school instruction to *"This is all well and good but I've never been in a real classroom that was taught this way,"* to *"I have so much material to cover, I don't know if I have time to do this."* These problematic issues were a good beginning for using the RTOP as a catalyst for inquiry into reformed teaching. The day ended when teachers were asked to make a short journal entry where they would reflect upon how their own teaching practice would score on the RTOP and how they felt about this. These journals were collected the next morning, and an hour long-focus group / discussion regarding RTOP was conducted, taped and transcribed.

For the next two days of the first week of the workshop, participants broke into small groups and worked through several selected activities from the modeling physics curriculum, led where appropriate by their master teacher instructors. Participants were invited to RTOP the master teachers as they taught. The invitation was issued to establish the importance of RTOP in providing reflective feedback of value to any and everyone interested in improving their teaching, including the instructors of the course. At times, the instructors reacted to ratings given by participants by saying things such as *"that rating is entirely too high – I didn't do that!"* Such reactions sparked deliberation and debate. In this way, both participants and instructors engaged in inquiry-based learning as a way of understanding reform. This practice was accompanied by nightly journal reflections upon reform.

1. "Dave and Ted's Excellent Adventure' (title taken from a journal entry)

Dave and Ted were part of the 4-person instructional staff. During the first week, they presented a model physics lesson in which participants were challenged to construct a clock consisting of a set of 10 balloons filled with water strung together with 10 meters of cord. When the linked balloons were hung from the top of a third floor building and then let go, the string of balloons were to strike the ground, creating 10 bursts of sound equidistant in time. In effect, participants were to construct a 10 interval clock on the

basis of their understanding of acceleration. This lesson is in fact a modern interpretation of an activity originally carried out and described by Galileo (Galilei, 1632).

- *My first thought as "a student" was COOL! It drew me in immediately. We were going to break stuff! One of the hardest parts to this activity, as a teacher, was to some how "forget" your previous knowledge in order to come at the problem "as a student." For me it was not that hard because my deep understanding of physics is equivalent to the kiddie pool. And that is the nice thing about this activity, is that it can help the students come to an understanding together . . . .Once back inside, the class came together to discuss the ways in which the teams arrived at their solution. This gave a chance to address any misconceptions any of us had about the lesson. I think that is high on the list of hardest objectives to reach: getting the students to come to terms with what they hold to be true. It is another to do this in a non-threatening way so that their students will learn from this. I feel Dave did a good job of doing this. I feel Dave gave a good example of modeling as a reformed teacher, he allowed the students to make the discovery on their own while facilitating the discussion afterwards without given us any preconceived ideas.*

- *In terms of the lesson design and implementation, the balloon drop served as a beautiful lesson for the following reasons. It respected students' prior knowledge by providing students with the opportunity to face their misconceptions through experimentation. It engaged students as members of a learning community, both small group and large group. Simple instructions about timing devices preceded the student exploration. Students began working almost immediately thereafter. To succeed, group members had to cooperatively build the idea, or hypothesis, and ultimately agree on method of testing. Students largely direct the lesson once it had begun.*

- *As I was a participant in Dave and Ted's (excellent adventure) balloon drop, I was enjoying the fact that I was a student. More specifically, I felt "empowered" by being able to explore on my own, most of the time without intervention. That was a very important lesson that I took away as a teacher from that activity. By being allowed to explore, I understood what it might be like to be a student in one of my classes. I am sure at many times it may have been difficult for Dave and Ted to* ***not*** *give hints or suggestions to the groups hypothesizing what was required to produce the desired impact sounds and "pace" of balloons dropping. I think Dave And Ted were successful in their goal of "student exploration" for that section of the lesson. . . . Overall, I enjoyed the fact that as students, we did most of the work and all of the exploration. I was proud of the ideas that we developed without the aid of a teacher.*

2. The "Ultrasonic Motion Detector" lab

The "Ultrasonic Motion Detector" lesson provided a prime opportunity to examine the notion of a "teachable moment" in new light. Because inquiry lessons respect and encourage student initiative, "teachable moments" are likely to emerge when least

expected. Making the most of them when they occur is not easy. This lesson provided participants with an opportunity to sink their teeth into the notion of a teachable moment and how such moments, even if they feature in inquiry teaching, can become hugely problematic.

- *As a student participating in the activity, I had mixed emotions. While the point about instantaneous acceleration was well taken, I personally thought that aspect of the lesson was far too drawn out. We went on for over 30 minutes about that one point. Granted, some of that time was devoted to philosophical differences (for example, teachable moments), but I had no problem accepting the obvious (to me) fact that objects can't attain a constant velocity from a dead stop without some acceleration. Apparently that was not obvious to some, so I felt like I had to wait around until they say the light./ What do you do about students who do catch on quickly (believe me, it was a fluke that I did this time!) --- how do you keep them actively engaged in the discussion? Not being a physics person, I was really bored with the discussion after the first five minutes. Enough about the bees already! . . . From the little experience I have had with white boarding, it seems to take an incredible amount of time, and time is something that is usually at a premium. I am already frustrated by the time constraints placed upon me by my curriculum (then add to that all the silly things we have to do) and I question how I am going to accomplish it all. I am interested in hearing how teachers maintain the "teacher as listener" position and still direct the discussion in the right direction without using large amounts of time. . . . Anyway, good job Ted! You've done a good job of modeling some aspects of reform teaching and stimulated my thing on the subject.*

- *The white board presentations started out well. But the teaching moment turned out to be a 30 minute lecture and resulted in being a teacher centered and directed activity. If the teacher would have held out for a few more white board presentations, the student discussion may well have emphasized the exact point the teacher made. The students were not given a chance to discuss and possibly come to that exact conclusion.*

Comment: The first participant was frustrated with the expenditure of 30 minutes on concepts which were obvious to her. Despite this, she was able to sense that other participants were likely in a situation where the patience of the instructor was enabling them to reach new realizations, and so she used the phrase "teachable moment". The incredible amount of time the instructor allowed to pass by dwelling on just one idea rendered problematic the notion of "teacher as listerner": when is too much listening a waste of time? And yet encouraging students to take control of the lesson is important in reform.

In the second reflection, the participant goes so far as to call the 30 minute excursion a "lecture". She interpreted the 30 minutes as being too exclusively guided by the instructor. If only he had allowed other students to enter by contributing their whiteboard ideas, *"the student discussion*

may well have emphasized the exact point the instructor made." Thus from the perspective of the first participant, the teacher was listening too much; from the perspective of the second participant, he should have listened even more (but more broadly). These two reflections point to the complexities, nuances and intricacies of the discourse that characterizes inquiry communities.

Week Two

With a week of modeling and "RTOPing" under their belt, it was now time for participants to embody their understandings of reform as revealed through the RTOP be designing a unit of their choice. For the first three days of the second week, participants broke into small groups self-selected by subject and grade level interests. Each group prepared a unit plan by assembling activities they had previously identified as outstanding and appropriate for their own students. Additionally, groups of teachers identified relevant state science and mathematics standards and attempted to refine their instruction by incorporatingreform teaching insights gained from both RTOP and the modeling curriculum. By topic, the resulting unit plans and activities produced were *ecological interactions, measurement, families of algebraic functions, Boyle's Law, chemical solutions, energy as a means of change, and bacterial growth* (SCI599, 2000).

On the last two days of the second week, chosen activities from these units were presented by each group in turn to the other course participants, some of whom acted as students working through the activities while others RTOPed the lesson. Instructors and the RTOP author also completed RTOP's on the lessons. The presentations were also videotaped for later analysis. Unit critiques, RTOP feedback and self-reflection were used to redesign the unit plans, which were placed upon the web for distribution (SCI599, 2000). Reflective journals kept by participants were collected and commented/responded to by a master teacher instructor.

An interesting development from the end of week two was a division amongst participants with regards to satisfaction in regards to how reformed their instruction actually was. Several participants felt the activities presented were not truly reformed, and that the RTOP evaluations wereapologetic – that is, participants were not rigorously enough in critiquing the degree of reform actually in evidence during instruction. After some debate, several participants revised lesson materials, and most participants decided to be more self-critical later in the course. In the final assessment, one participant described this as the lowest point in the course...

Um.. lets see, Ah.. I think the low point was the first set of ah.. lesson plans that were not as in [hesitation] creative as I think we could have done...

Week Three

The first half day of week three was spend in a group focus group/discussion on the general strengths and weaknesses of the unit plans with an overview of the philosophy underlying the presented lessons. This was audio recorded and transcribed.

Fueled with new insights into reformed teaching garnered from the previous week's activity and the reflection upon these activities, the remainder of the first two days of the final week was spent operating in new groups, developing unit plans on new topics. Some groups chose to spend additional time enlarging and refining previous work, but most developed new activities including *atmospheric pressure, chemical bonding, the periodic table, mass conservation, direction and displacement, population dynamics* (extended and refined from *bacterial growth), sound and music, cellular basis of activity* and *water quality* (SCI599, 2000). The third and fourth day of week three were spent again presenting, critiquing and videotaping selected lesson presentations from the newly assembled units. Again, reflections were kept in a journal, and read and commented upon / responded to in writing by a master teacher instructor. Participants revised the units into a final form for the website, and this was also copied to a CD-ROM later mailed to all participants. On the final day of the course, an open ended questionnaire was collected from all participants, and a final course wrap-up / focus group was held. The final questionnaire forms the primary data source for this article, and the final focus group was taped and transcribed as well.

On the last day of the final week, focus group interviews were held with participants to gather their thoughts on various aspects of the course. One of the questions asked students to identify the high and the low points of the course. A couple of the students made remarks that reveal rather succinctly how the RTOP had helped transform their understanding of their own teaching practice.

- *The RTOP has changed how I will teach science from now on. Um.. it has identified for me some.. some really fairly simple adjustments that I can make that I think will be extremely helpful for my students and has helped to identify for me uh.. some of the frustrations that I have felt when teaching my students. You know... knowing that I myself had a model that I was following but never actually getting that clearly across to my students and realizing that.. that is a problem. So now that I have identified it, I can deal with it and I find that it is not as hard to deal with as you might think it just takes a little practice.*

- *Um.. lets see, Ah.. the high point was really taking the RTOP and getting ideas for how to transform the lesson plan, like keep the lesson plan you already wrote, keeping it RTOP. What we can incorporate that we didn't already do and also just having that solid reference for how.. how to improve um.. my lessons. That I know that if I get a high RTOP score, my students will, you know have better conceptual understanding. So its very easy for me to go through this list to improve. I also like the experience of friends and I really appreciate carrying everyone's ideas here.*

Comment: These two entries express two different ways the RTOP has guided self-reflected change. For the first participant, the RTOP has become a frame of reference that allows her to understanding her own teaching with sufficient clarity that she can communicate what she is doing so that her students experience the authenticity of her teaching. On the other hand, the second participant is RTOP has become a tool to improve her lessons allowing her to transform existing materials into reformed lessons. From our point of view, both approaches allow participants to continue using the RTOP for self-reflective change.

Final focus group interviews also made interesting juxtaposition with initial interviews held after the end of RTOP training on Day 2. At that time, several participants expressed extreme disbelief with regards to the reforms espoused by RTOP:

...Yeah, just drop and say I ain't doing this RTOP thing... just an example; this kid was making fun of this one kid so I sent him down to take the attendance, and said look you guys you can't keep doing this, this is not right to do to somebody, O.K. how would you like it if we did this to you. O.K. were going to talk to about, make something up to someone about... It's not an RTOP lesson, and that comes up so many times in the middle school, you know.

...I'll agree with that, that it is, there are certain in a traditional mathematics curriculum the way its sequenced traditionally.. it.. there are something's that lend them selves wonderfully to their investigative constructavist approach and there are something's that in first place don't, but the longer.. the more experience you have in doing it the more you learn to look at specific ideas and topics, and concepts from.. from.. the tools and the back ground of how do I get the kids to investigate this before I show them step one, step two, step three, here's how you get the answer and.. think about what the answer means now.

Other participants expressed fear of the assessment of their teaching practices by classroom visitors, particularly from school administrators who they felt would not be able to appropriatly use the instrument. Participants in general felt that RTOP was most appropriate for peer mentoring.

...lets say that the students came up with what you were going to do but that's not the day I was evaluated, the next day the evaluator came in and were doing what the students and I'm going to get a zero on this day, but it will certainly student directed, but there no evidence, there's no evidence that they came up with the idea which is one of the particular indicators, that they actually initiate the idea. So that's not always obvious on a daily basis.

...when you do that it becomes a top down issue, you know it comes and (gets you?) the... readminstration and when you do a peer-peer... When its peer-peer it is such, it is taken as something as let me help you, let us help each other. You know

its a sharing sort of difference opposed to the administrative instrument they use as a check...

So it just seems like everybody's talking a lot about, and were worried other people evaluating us and like being I mean, it never feels good to be judged right, and this whole thing is starting to sound like really judgmental and were worried about the word reform. It sounds like a judgmental word, or what ever. But the whole, if we could just forget all that for a minute and be like, the whole thing is like you suck. If it actually improves, the students scores on these conceptual tests it, we think that's a really good thing, you know, we can look at this and purposely, you know, try to include more of the elements that are on teaching, and I found it actually really exciting because I been kinda, you know, I mean in my student teaching I was really frustrated with like thinking I need three more hours a day to lesson plan and...

One participant made an interesting comment describing her own previous classroom assessments as mystical ...

Except for a lot of comments that good things weren't happening, you know there was a lot of people who were extremely mystical that there were, that she wasn't at all a (good) teacher, just give her... you know that's comments that came out yesterday, and so..

Participant journals also showed remarkable changes in language use. While early reflections contained many general terms and concepts such as "stuff", "active", "hands-on" and "fun", participants were now using words such as "students' prior knowledge", *"face their misconceptions trough experimentation", "engaged students as members of a learning community", "agree on method of testing", "student exploration" and the like.*

Data Analysis & Results

Volumes of data had been collected but how did they inform the question underlying the study: Was it feasible to use the RTOP as a catalyst for self-reflective change?" Researchers reflected on the data for the purpose of coming up with ways in which "feasibility" manifested. Three major assertions based upon catalytic characteristics emerged from the collected data and consensus formed around the generalized assertions categorized in Table 1.

Table 1
Assertions and Catalytic Characteristics of RTOP Emerging From This Study

Assertion One: RTOP is strongly credible, authentic, justifiable, and believable instrument for supporting reform as reported by teachers, and has strong research and face validity as described by educational researchers.

1.1 Use of the instrument as a catalyst for self-reflective change fostered in the group the very properties the instrument measures. The class itself was a standards-based inquiry community. More specifically, the graduate course itself scored high on the instrument.
1.2 The instrument has a history of use in science and mathematics education reform that provided strong initial credibility in the eyes of participants.
1.3 The range of the instrument more than spanned the reform practices of all participants including staff. (There is no ceiling effect).

Assertion Two: RTOP provided an open, observable, comfortable and empowering guide catalyzing reform teaching.

2.1 Instructors of the course bought into the catalytic use of the instrument.
2.2 The use of the instrument in assessing lessons was itself an invitation into inquiry.
2.3 The instrument was openly used by all members of the group, including staff, to catalyze self-reflective change.
2.4 The instrument was openly used by both staff and students to assess teaching of both staff and students – it is used by everyone to assess the teaching of anyone. The community was very open in regard to any feedback that enhances self-reflective change.

Assertion Three: Evidence exists to support claims for individual and community change in terms of language, reconceptualization, self-critique. Evidence exists that reform was sustained in and beyond the course.

3.1 The use of the instrument supports and catalyzes the development of language particularly suited to assist in conceptual development of standards-based inquiry.
3.2 The use of the instrument allowed students to become aware and articulate changes about their own teaching.
3.3 The use of the instrument supported and catalyzed the critique of standards-based inquiry teaching.
3.4 The use of the instrument catalyzed a multitude of ways in which participants could reconstrue their understanding of their own teaching and develop specific ways to direct and redirect their own development.
3.5 RTOP scores for teaching occurring during the latter half of the class (55) would be substantially higher than RTOP scores for teaching during the first half (35).
3.6 The use of the instrument catalyzed the organization of other inquiry venues that sustain the self-reflective change after the course is over.
3.7 Participants continued to use the instrument to reflect upon their teaching
3.8 Participants were observed to have changed their teaching/curriculum when they return to their classrooms. (eg. Dave T and SCI 351/352)

Conclusion

We suggest that the quantitative and narrative evidence provided in this paper together with the interpretive commentary support the catalytic use of a classroom observation protocol to stimulate self-reflective change toward standards-based inquiry in science and mathematics classrooms. As might be expected, participants have found their own ways of bringing the RTOP into their movement toward reform. Some have chosen to keep the RTOP on the periphery; others have given it the old college try. Many of the participants in the study have volunteered to continue with the RTOP as a reflective device in a follow-up project that will be ongoing for the next few years. This will enable the researchers to further monitor and support self-reflective change as further research into reformed professional development is undertaken. While other rubrics or matrices are available for framing reformed teaching, we believe the RTOP offers special advantages because of its established psychometric properties as a quantitative measurement instrument (Piburn et. a., 2000). Because of its very strong validity in predicting achievement gains in science and mathematics in large reform projects (Lawson, et. al., in press), in the eyes of practicing teachers, it possesses a degree of credibility. The openness of staff to have their own teaching "exposed" to RTOP dramatically increased both the credibility of the instrument, the credibility of the staff, the credibility of the course, and the credibility of the research. At the same time this use of RTOP expanded the willingness of participants to have their teaching assessed using RTOP. With increasing use the instrument contributed to an evolving pedagogical vocabulary of participants. More significantly, the language in the instrument provided

participants with specific concepts and terms for thinking about and talking about reform. Its repeated use led to what might be called a "discourse of reform" which both supported and helped constitute the emerging inquiry community.

We believe the research reported here, together with the interpretive commentary, establishes the feasibility of using RTOP as a catalyst for self-reflected change.

References

AAAS (1989). Project 2061: *Science for All Americans: A Project 2061 Report on Literacy Goals in Science, Mathematics, and Technology*. Washington, D.C.: American Association for the Advancement of Science.

Adams, P. & Krockover, G.H. (1997). *Stimulating professional development through use of interviews and observations*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Oak Brook IL, March 21-24, 1977. ED405 212.

Adams, P. & Krockover, G.H. (1999). Stimulating constructivist teaching styles through use of an observational rubric. *Journal of Research in Science Teaching, 36(8)*, 955-971.

Arizona Board of Regents (2000). ABOR / Eisenhower mathematics and science education act request for proposals. Phoenix AZ: ABOR. Available from http://www.abor.asu.edu (Jan 2000).

Arizona Teacher's Excellence Coalition (AzTEC)

Calhoun, D.O., Bohlin, C., Bohlin, R. & Tracz, S. (1997). *The mathematics reform movement: Assessing the degree of reform in secondary mathematics classrooms*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, March 24-28, 1997.

Forman, E.A. (1996). *Mathematics and talk: Evaluating discourse in reform classrooms*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY, April 8-12, 1996.

Galilei, G. (1917). *Dialogues concerning two new sciences*. Orig 1632 trans by Favaro, A. NY: Dover.

Gallagher, J. & Parker, J. (1995). *Secondary Science Teacher Analysis Matrix (SSTAM)*. East Lansing, MI: Michigan State University.

Hallouin, I.A. & Hestenes, D. (1987). Modeling instruction in mechanics. *American Journal of Physics, 55, 455-462*.

Hestenes, D. (1987). Toward a modeling theory of physics instruction, *American Journal of Physics, 55: 440-454*.

Horizon Research Inc. (1998). *1997-98 Local Systemic Change Revised Classroom Observation Protocol*. Chapel Hill: NC, The Author.

Lai, M.K. & Young, D.B. (1997). *(Science) education standards yield standard data-collection instruments.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, March 24-28, 1997.

Lawson, A. E. (1995). Science teaching and the development of thinking. Belmont, CA: Wadsworth Publishing. Pp. 206-208.

Lipsey, M.W. (1994). Theory as method: Small theories of treatments. In L.B. Sechrest & A.G. Scott (Eds.), Understanding causes and generalizing about them. *New Directions for Program Evaluation, 63 (Fall),* 5-25.

MacIsaac, D.L. & Cole, D. (1999). *Arizona standards-based modeling for middle and high school science and mathematics teachers: A proposal for the 2000 Eisenhower awards.* Flagstaff, AZ: The Author.

Mervis, J. (1998). Mixed grades for NSF's bold reform of statewide education. *Science, 282(4),* 1800-1805.

Middleton, J. A., Horn, P., Ridley, D. S., Luft, J. D., Wyckoff, S., Brush, T., MacIsaac, D., & Bloom, J. (1999). AZTEC-Arizona Teacher Excellence Coalition: A Statewide Partnership. A proposal to the US Department of Education Partnership Grants for Improving Teacher Quality Program. Tempe, AZ: Arizona State University. Available from http://purcell.phy.nau.edu/aztec (March 2001).

NCTM (1989). Curriculum and Evaluation Standards for School Mathematics. Reston, VA: The author.

NCTM (1991). Professional Standards for Teaching Mathematics. Reston, VA: The author.

NCTM (1995). Assessment Standards for School Mathematics. Reston, VA: The author.

NCTM (2000). Principles and Standards for School Mathematics. Reston, VA: The author.

NRC (1996). National Science Education Standards. Washington, D.C.: National Academy Press.

NRC (2000). Inquiry and the National Science Education Standards. Washington, D.C.: National Academy Press.

O'Sullivan, R.G. (Ed.). (1995). Emerging roles of evaluation in science education reform. *New Directions for Program Evaluation, 65 (spring).*

Piburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). *Reformed Teaching Observation Protocol (RTOP): Reference Manual.* (ACEPT Technical Report No. IN00-3). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Porter, A.C. (1995). *Standard setting and the reform of high school mathematics and science.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 18-22, 1995.

QUASAR. (1992). *QUASAR Documentation: Classroom Observation Instrument*. Pittsburgh, PA: QUASAR Project, Edward A. Silver, Director, Learning, Research and Development Center, University of Pittsburgh.

Piburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). *Reformed Teaching Observation Protocol (RTOP)Reference Manual* (ACEPT Technical Report No. IN00-3). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Priestly, H, Priestly, W.J., Sutman, F.X., Schmuckler, J.S., Hilosky, A. & White, M. (1998). Evaluating the use of the inquiry matrix. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, San Diego, CA, April 19-22, 1998.

Sawada, D. (1999). *The Development of the Reformed Teaching Observation Protocol (RTOP).* Technical Report No. IN99-1). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sawada, D. (1999). *Psychometric Properties of RTOP.* Technical Report No. IN99-2). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reformed Teaching Observation Protocol (RTOP)* (ACEPT Technical Report No. IN00-1). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sawada, D. Piburn, M., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). *Reformed Teaching Observation Protocol (RTOP) Training Guide* (ACEPT Technical Report No. IN00-2). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

SCI599 Course WebPages (2000). Flagstaff, AZ: NAU Department of Physics & Astronomy. Available from http://purcell.phy.nau.edu/courses/00/summer/SCI599/ (March, 2001).

Silver, E.A. & Cai, J. (1993). *Schemes for analyzing student responses to QUASAR's performance assessments: Blending cognitive and psychometric considerations.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA., April 13, 1993.

Tittle, C.K. & Pape, S. (1996). *A framework and classification of procedures for use in evaluation of mathematics and science teaching.* Paper presented at the annual meeting of the American Educational Research Association, New York, NY., April 8-12, 1996.

Tomoff, J. (2000). *Measuring NCTM-recommended practices and student achievement with TIMSS.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 28-28, 2000.

Weir, D. (2000). *Connecting the standards to achievement: Mathematics classroom practices and student outcomes.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 28-28, 2000.

Wells, M, Hestenes, D. & Swackhamer, G. (1995). A modeling method for high school physics instruction, *American Journal of Physics,* 63, 606-619.

Young, J.J., Brett, B., Squires, S. & Lemire, N. (1995). *Symposium: A new observation tool for looking at inquiry-based teaching and learning.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 18-22, 1995.